

Random Tree Model of Meaningful Memory

Weishun Zhong,¹ Tankut Can^{1,2}, Antonis Georgiou,^{1,3} Ilya Shnayderman^{1,3},
Mikhail Katkov^{1,3} and Misha Tsodyks^{1,3,*}

¹*School of Natural Sciences, Institute for Advanced Study, Princeton, New Jersey 08540, USA*

²*Department of Physics, Emory University, Atlanta, Georgia 30322, USA*

³*Department of Brain Sciences, Weizmann Institute of Science, Rehovot 76100, Israel*



(Received 11 December 2024; accepted 30 April 2025; published 13 June 2025)

Traditional studies of memory for meaningful narratives focus on specific stories and their semantic structures but do not address common quantitative features of recall across different narratives. We introduce a statistical ensemble of random trees to represent narratives as hierarchies of key points, where each node is a compressed representation of its descendant leaves, which are the original narrative segments. Recall from this hierarchical representation is constrained by working memory capacity. Our analytical solution aligns with observations from large-scale narrative recall experiments. Specifically, our model explains that (1) average recall length increases sublinearly with narrative length and (2) individuals summarize increasingly longer narrative segments in each recall sentence. Additionally, the theory predicts that for sufficiently long narratives, a universal, scale-invariant limit emerges, where the fraction of a narrative summarized by a single recall sentence follows a distribution independent of narrative length.

DOI: [10.1103/physrevlett.134.237402](https://doi.org/10.1103/physrevlett.134.237402)

Narratives are a naturalistic form of stimuli for examining the structure and organization of human cognition [1–6]. Bartlett's classic work [7] established a link between narrative memory and comprehension, noting significant individual variations in both. Later studies highlighted that narrative comprehension and recall are deeply connected with semantic knowledge and schemas, which aid in interpreting new information using prior knowledge [8–11]. It appears therefore that the complexity of meaningful memory cannot be captured by simple models with a few postulates. In contrast, psychological studies often relate memory to linguistic organization without predicting quantitative recall features [12,13], unlike studies of random materials, which yield numerous quantitative results [14]; for example, some of the authors developed a model predicting the relationship between recalled items and retained items, well supported by experiments [15]. Unlike random lists, narratives have complex structures across abstraction levels, and recall is not just about word count since people do not recall meaningful material verbatim. Recall length alone may not indicate quality, as a concise summary can better convey meaning than a longer, unstructured recall. Thus, a relevant theory must address these complexities. The model presented

below is inspired by recent large-scale narrative experiments [16]. Unsurprisingly, we observed that recall is qualitatively different from that of random lists: People tend to recall information from the narrative in the same order as it is presented [16,17], as opposed to much more variable recall order of random lists [18–20]. For longer narratives, people compress progressively larger pieces of information from the narrative into single sentences, resulting in a sublinear growth of recall length with the narrative size [16].

In the current study, we show that despite the more complex nature of meaningful recall, the above statistical regularities can be captured by mathematical modeling. To this end, we propose a simplified model of narrative recall that focuses on describing the statistical aspects of recall across large populations of people rather than addressing individual differences between people and narratives.

Our model is based on two basic principles. The first principle describes the memory encoding of a narrative. Following multiple previous studies, we consider a narrative as a linear sequence of clauses, i.e., the shortest meaningful pieces of text (see, e.g., [21]). We assume, however, that after people comprehend the narrative, they form a treelike memory representation of it where, at each level of the tree, information units of a given level of abstraction are encoded [22–24], with each unit representing a summary of a specific segment of the narrative. In particular, the root of the tree represents the whole narrative; the upper-level nodes represent pieces of the narrative that correspond to its main key points [25], and lower levels represent shorter pieces corresponding to progressively finer points [26–29]. Moreover, we assume

*Contact author: mtsodyks@gmail.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

that a piece corresponding to a particular key point is split into pieces corresponding to the finer points that elaborate the original key point, resulting in a tree structure of a narrative. We see the evidence for this type of memory representation in the fact that people can easily summarize a familiar narrative at different levels of detail, which we believe is achieved by retrieving the node representations at different tree levels; e.g., retrieving upper nodes evokes the most abstract memory of a narrative in terms of its most general points. An implicit assumption made here is that continuous stimuli, such as a text, are represented by a discrete set of nodes. This is also supported by experiments that show humans naturally segment sequential stimuli (e.g., movies, personal experiences) into discrete units [30–35]. There is also extensive support for the claim that narratives, from folktales [36] to oral narratives [21,37], have such a hierarchical, treelike structure. We argue that this hierarchy is reflected in the mental representation. Finally, as we observed in our previous publication [16], people only remember a fraction of the narrative clauses as attested by recognition experiments; in addition, it seems plausible that not all recognizable clauses may be integrated into the tree representation. We therefore leave the size of the tree N as a subject-dependent parameter that should be estimated from the data, as we show below.

The second principle reflects the way people recall a given narrative based on its memory representation described above. We assume that people mentally traverse the tree from upper to lower levels, keeping the intermediate nodes in working memory, to constantly maintain the integrity of the narrative in mind. It is in this respect that we believe the recall of meaningful narratives is fundamentally different from the recall of random lists of words. In particular, working memory capacity limits the number of nodes that can simultaneously be kept in mind, thus limiting the number of levels that can be reached during recall. We further assume that when reporting a particular node during recall, people use a single clause to express the content of the corresponding piece.

It seems plausible that a specific instantiation of the memory tree depends both on the narrative structure and on the way a particular person comprehends and subsequently memorizes it. Hence, in order to translate the above two principles into a quantitatively constrained model, we make a highly simplifying assumption about the statistical ensemble of trees encoding different narratives of a particular length in the minds of different people. We build this ensemble by a recursive self-similar process in which a narrative is split into up to K pieces by the random placement of $K - 1$ boundaries, and each piece is similarly split into up to K pieces, and so on. We begin this process from the whole narrative and continue it recursively from one layer to the next until splitting stops (see details in Supplemental Material Sec. A [38]). The value of K could be considered a free parameter to be determined from the

results. We assume however that K is related to working memory capacity. Indeed, people can effortlessly summarize a narrative at the highest level of abstraction, which in our scheme would correspond to activating the representations of the upper-level key points. Therefore, we assume that narrative representations are constructed in such a way that summarizing the whole narrative (or any of its meaningful segments) does not overwhelm working memory capacity and therefore chose $K = 4$ [42].

We then illustrate the recall process playing out on this tree, assuming that people systematically traverse the memory representation by following each route beginning from the root, until they either reach a leaf or a lowest reachable (D th) level [the levels above the gray shaded area in Fig. 1(a)]. We again consider $D = 4$ because it corresponds to working memory capacity. The retrieval trajectory is shown in red on Fig. 1(a). One can see that some of the nodes reached by a retrieval trajectory contain single narrative clauses, while others contain multiple clauses (up to 4 for the illustrated tree). Since we assume that each node, no matter its level, is recalled with a single recall clause, the length of recall is defined by the number of retrieved nodes, while each recall clause summarizes the corresponding number of narrative clauses. It is important to emphasize that while the ensemble of memory representations described above involves a random set of trees, the hypothesized recall process itself is deterministic; i.e., the number of recall clauses and the pieces of the narrative that each one of them corresponds to are fully determined by the memory representation of the narrative.

The model described above can be simulated numerically and solved mathematically, as we show below. The length of recall (the number of clauses in recall C) averaged over the ensemble of trees with N clauses exhibits a characteristic saturating form for increasing N ; see Fig. 1(b) for this curve computed with $K = D = 4$. Other choices of these parameters are illustrated in Supplemental Material Sec. G [38]. We also computed the distribution of the number of narrative clauses summarized by single recall clauses (i.e., compression ratios of different recall clauses) shown in Fig. 1(c) for compression ratios up to 10 and for different N . One can see that, as expected, for longer narratives the compression ratios of recall clauses tend to increase.

To obtain the analytical approximations for these numerical results, we derive a recurrence relation for the distribution of the number of clauses represented by nodes at a given level l , $P(n^{(l)})$, with

$$P(n^{(1)}) = \delta(n^{(1)} - N), \quad (1)$$

since the upper node represents all of the narrative clauses retained in memory after acquisition. When a node at level l with $n^{(l)}$ clauses is split randomly into K nodes of level $l + 1$ (some of them empty), the distribution of the size of each of the resulting nodes can be computed by the “stars

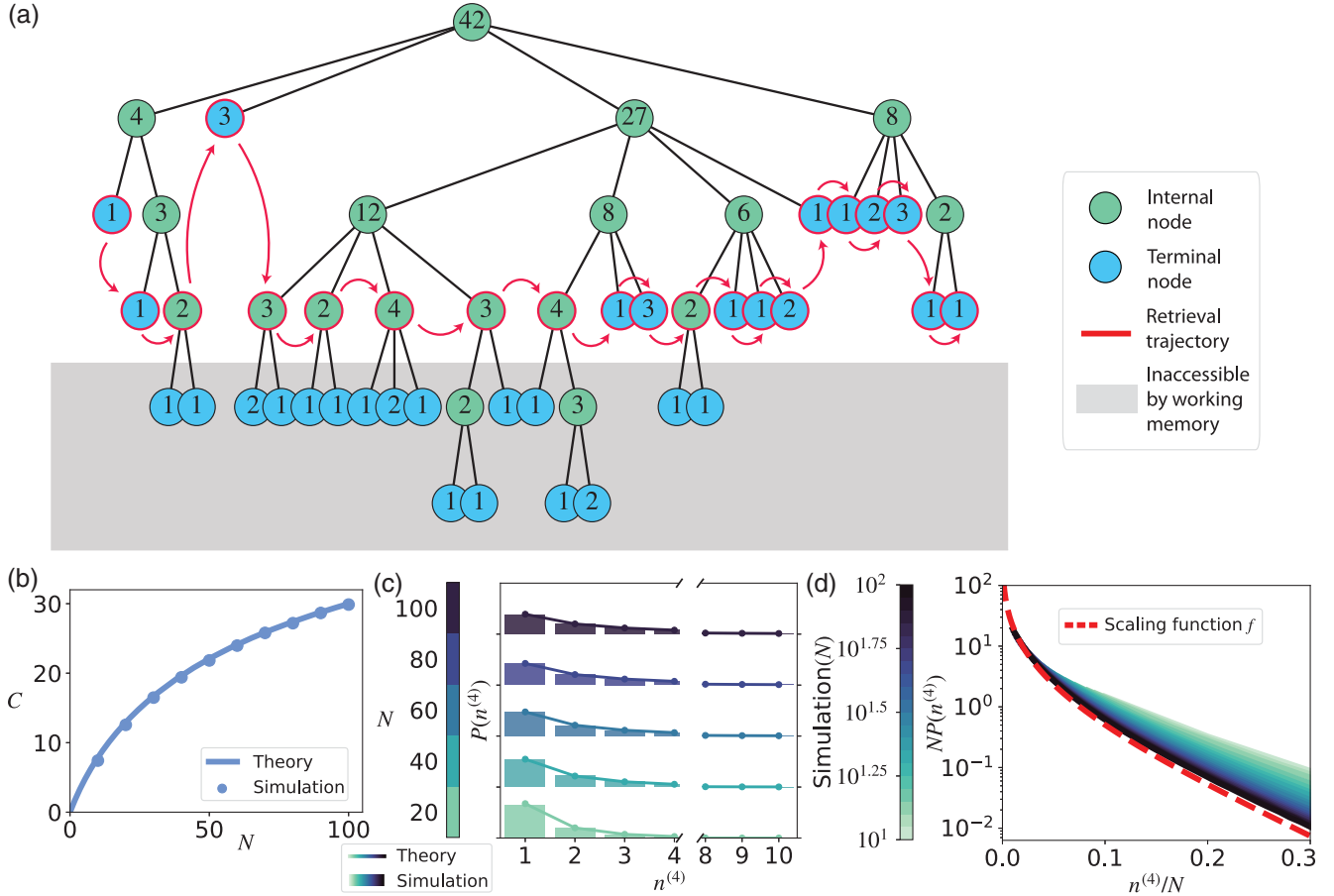


FIG. 1. Ensemble of random trees. (a) Schematics of memory retrieval from a random hierarchical representation. An example of a single realization of the random tree created by the model for $N = 42$ encoded clauses with branching ratio $K = 4$ (empty nodes are not shown). Internal nodes are shown in green, while the terminal nodes (leaves) are shown in blue. The tree generating process starts with the whole narrative contained in the root node (level 1), which is subsequently split into up to four chunks at the next level. The splitting continues self-similarly until either a chunk fails to split further (e.g., the blue “3” at level 3) or becomes a single clause (the blue “1s”). The gray shaded area illustrates the limit imposed by working memory capacity as retrieval starts by descending from the root (retrieved nodes are shown in red). (b),(c) Comparison between analytical solution and numerical simulations. (b) Mean recalled length C as a function of the encoded length N . Numerical simulations are averaged over 10^4 realizations of random trees (see details in Supplemental Material Sec. A [38]). (c) Distribution of the chunk size at the D th level $n^{(D)}$ ($D = 4$) given root size $n^{(1)} = N$; the range between tick marks in the y axis corresponds to $[0, 1]$. (d) Distribution of compression ratios scaled by N as a function of the compression ratios divided by N . Simulations of different N 's are shown in different shades of green. The red dashed line is the asymptotic scaling function from Eq. (7).

and bars” method [43] by considering all possible configurations of $n^{(l)}$ stars and $K - 1$ bars, with a fixed number $n^{(l+1)}$ of stars to the left of the leftmost bar. This results in the following recursive expression (for $n^{(l+1)} \leq n^{(l)}$, 0 otherwise):

$$P(n^{(l+1)}|n^{(l)}) = \frac{Z_{K-1}(n^{(l)} - n^{(l+1)})}{Z_K(n^{(l)})}, \quad (2)$$

$$Z_K(n) := \binom{n + K - 1}{K - 1}. \quad (3)$$

Applying this recursive expression and summing over intermediate levels results in the distribution of the node sizes in level D ,

$$P(n^{(D)}) = \sum_{\substack{0 \leq n^{(D-1)} \\ \leq \dots \leq n^{(1)}}} \prod_{l=1}^{D-1} P(n^{(l+1)}|n^{(l)}) P(n^{(1)}). \quad (4)$$

Note that Eq. (4) defines a Markov chain; see Supplemental Material Sec. B [38] for more details. In this version of the stars and bars model, some of the nodes may end up being empty (if two or more bars are adjacent), and the

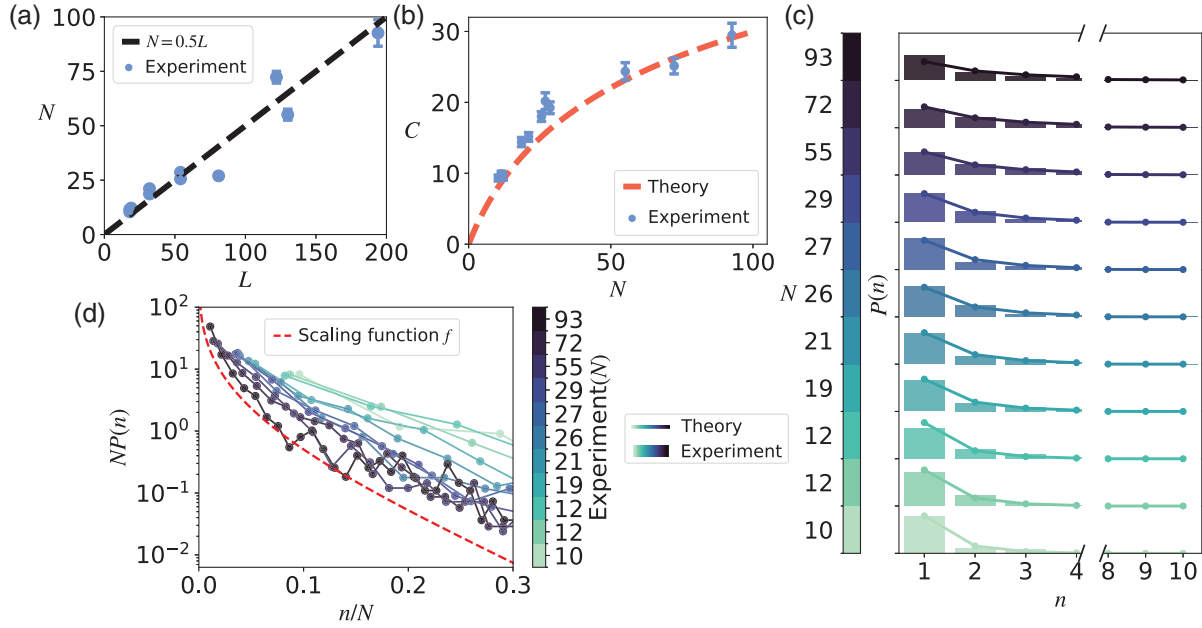


FIG. 2. Comparison between theory and experiment. (a) Average size of the tree memory representation of each narrative (N) estimated as explained in the text, plotted vs narrative length (L) for 11 narratives in the dataset. Dashed line corresponds to $N = 0.5L$. (b) The mean number of recalled clauses C vs average N for all 11 narratives. Blue filled circles are data. Red dashed line is the theoretical prediction obtained from Eqs. (4) and (5) with $K = D = 4$. Error bars in (a) and (b) are the standard error of the mean. (c) Normalized empirical histograms of compression ratios for all subjects separately for each narrative, as measured from mapping recalled clauses back to the narrative clauses. Data for different narratives are shown in color corresponding to the color bar marked with values of N for each narrative. Solid lines are theoretical predictions obtained from Eq. (4) with $K = D = 4$. Range between tick marks in the y axis is $[0, 1]$. (d) The distribution of experimentally measured compression ratios relative to N approaches the universal scale-invariant scaling function f in Eq. (7) as N increases.

average number of nonempty nodes at level D will define the average length of recall $C = K^{D-1}[1 - P(n^{(D)} = 0)]$, which is found to be (see Supplemental Material Sec. B [38] for derivations and its connection with the Riemann zeta function)

$$C = K^{D-1} \sum_{m=1}^N \binom{N}{m} \frac{(-1)^{m+1}}{Z_K(m)^{D-1}}, \quad (5)$$

while the renormalized distribution of nonzero node sizes corresponds to the distribution of compression ratios introduced above. As N increases, the probability to have an empty node approaches zero, and this equation therefore predicts that C saturates at K^{D-1} (i.e., 64 for our choice of $K = D = 4$) for long narratives. For $K = D = 4$, the analytical results obtained with Eqs. (4) and (5) are very close to numerical simulations, as shown in Figs. 1(b) and 1(c). In particular, one can see from Fig. 1(c) that as the narrative gets longer, the distribution of compression ratios shifts to the right; i.e., recall clauses tend to summarize progressively bigger pieces of the narrative.

For long narratives, the distribution of compression ratios normalized by narrative size ($s = n/N$) asymptotes to a universal scaling function $f(s)$ (see derivations in Supplemental Material Sec. B [38]):

$$P(n^{(D)}) = \frac{1}{N} f(n^{(D)}/N), \quad (6)$$

$$f(s_D) = \prod_{l=1}^{D-1} \int_{s_{l+1}}^1 ds_l \rho(s_{l+1}|s_l) \rho(s_1|1), \quad (7)$$

$$\rho(s_{l+1}|s_l) := \frac{K-1}{s_l} \left(1 - \frac{s_{l+1}}{s_l}\right)^{K-2}. \quad (8)$$

In this limit, $f(s)$ is a scale-invariant probability density function for the normalized compression ratio s , which in particular implies that the average compression ratios grow proportionally to the narrative length. The shape of the scaling function computed with Eq. (7) with $K = D = 4$ is shown in Fig. 1(d) in red, together with finite- N simulations from the random tree model.

We compare our theoretical predictions with experimental data from 11 narratives, ranging from 19 to 194 clauses, including eight from our previous study [16] and three selected for this research [44]. Most narratives were either from Labov's study [21] or generated by GPT-4 [45] using Labov's templates, with edits for clarity and spelling. Two narratives were sourced from [46]. We used Labov's clause segmentation, applying it to GPT-4-generated texts as well. Each narrative was presented to 100 subjects via Prolific [47],

who were instructed to recall them as accurately as possible. Following the protocol in [16], we used GPT-4 to segment recalls into clauses, with identical prompts (see Supplemental Material Sec. C [38]). To compute the compression ratios, we used GPT-4 to map each recalled clause to its corresponding segment in the original narrative (see Supplemental Material Sec. D [38]). For each recall clause, we identified the set of narrative clauses it mapped onto; this set constituted the narrative segment summarized by the recall clause, and its length defined the compression ratio for that clause. Essential to aligning these results with theoretical predictions is estimating the number of clauses in each subject's memory representation (N). We calculated this as the total number of narrative clauses mapped by at least one recall clause, assuming unmapped clauses were forgotten or not integrated into the tree. The average N for each narrative is approximately half its length, as shown in Fig. 2(a). With this assumption, average recall lengths and compression ratios up to 10 align well with theoretical predictions in Figs. 2(b) and 2(c) without tuning parameters. Figure 2(d) illustrates that for longer narratives, the compression ratio distribution relative to N approaches the scale-invariant form predicted by Eqs. (6) and (7), at a rate similar to corresponding simulations in the same range of N 's shown in Fig. 1(d).

Using compression ratios to quantitatively characterize recall is intuitively appealing but presents challenges. It involves mapping individual recalls to original narratives, which is subjective and time consuming, requiring advanced language models such as GPT-4 [45]. Random checks show that humans agree on clauses with small compression ratios (more concrete ones) but often disagree on those with large compression ratios (more abstract ones). This might be because mapping abstract clauses requires understanding both the narrative and the recall, which according to our model implies constructing corresponding tree representations and matching them to each other, both of which could vary between individuals. To test the robustness of mappings performed by language models, we used the DeepSeek model [48]. Clause-by-clause comparisons between GPT-4 and DeepSeek revealed inconsistencies, especially for longer narratives and recall clauses with high compression ratios. Despite this variability, the statistical features of recall—the dependence of the average recall length on estimated size of the tree representations and compression ratio distributions—align with our theory's predictions when using mappings from both models (see Supplemental Material Sec. F [38] for details).

In summary, our study demonstrates that despite the complexity of narrative recall, certain statistical trends can be modeled using two key principles: random treelike memory representations and a deterministic recall process, both constrained by working memory. The model accurately predicts the average recall length and the distribution of compression ratios for recall clauses based on narrative tree sizes, which we estimated individually for each subject

using large language model-performed mappings. We found that the average narrative tree size was about half the total number of clauses. This proportion is lower than the 70% of clauses retained in memory from recognition experiments [16], suggesting some retained clauses are not integrated into the tree. Our model does not address how memory representations form during the initial acquisition, which warrants future study, as does narrative production, which in our view involves recalling autobiographical events while also adding other information needed to make the recall understandable. Our narratives were of a single type, and future research should explore whether our model applies to other types, like fictional narratives.

The model's two parameters—the maximal splitting ratio in the tree and the maximal recall depth—align with human working memory capacity and were not fine-tuned to the data. This highlights working memory's critical role in narrative recall, a hypothesis for further exploration. Notably, we observed significant variability in individual recalls, potentially reflecting differences in working memory capacity. Indeed, most recall lengths fall between model predictions corresponding to capacities of two and six (see Supplemental Material Sec. G [38]), matching the reported range of working memory capacities [42].

Acknowledgments—W. Z. is supported by Eric and Wendy Schmidt Member in Biology Fund and the Simons Foundation at the Institute for Advanced Study. A. G. is supported by the Martin A. and Helen Chooljian Member in Biology Fund and the Charles L. Brown Member in Biology Fund. M. K. is supported in part by a grant from Fran Morris Rosman and Richard Rosman. M. T. is supported by the Simons Foundation, MBZUAI-WIS Joint Program for Artificial Intelligence Research and Foundation Adelis.

Data availability—The data that support the findings of this Letter are openly available [44].

-
- [1] Perry W. Thorndyke, Cognitive structures in comprehension and memory of narrative discourse, *Cogn. Psychol.* **9**, 77 (1977).
 - [2] Gordon H. Bower, Experiments on story understanding and recall, *Q. J. Exp. Psychol.* **28**, 511 (1976).
 - [3] Hongmi Lee, Buddhika Bellana, and Janice Chen, What can narratives tell us about the neural bases of human memory?, *Curr. Opin. Behav. Sci.* **32**, 111 (2020).
 - [4] Uri Hasson, Janice Chen, and Christopher J Honey, Hierarchical process memory: Memory as an integral component of information processing, *Trends Cognit. Sci.* **19**, 304 (2015).
 - [5] Yulia Lerner, Christopher J Honey, Lauren J Silbert, and Uri Hasson, Topographic mapping of a hierarchy of temporal receptive windows using a narrated story, *J. Neurosci.* **31**, 2906 (2011).

- [6] Lucas M Bietti, Otilie Tilston, and Adrian Bangerter, Storytelling as adaptive collective sensemaking, *Top. Cognit. Sci.* **11**, 710 (2019).
- [7] Frederic Charles Bartlett, *Remembering: A Study in Experimental and Social Psychology* (Cambridge University Press, Cambridge, England, 1995).
- [8] Joseph W. Alba and Lynn Hasher, Is memory schematic?, *Psychol. Bull.* **93**, 203 (1983).
- [9] Christopher Baldassano, Uri Hasson, and Kenneth A. Norman, Representation of real-world event schemas during narrative perception, *J. Neurosci.* **38**, 9689 (2018).
- [10] David E. Rumelhart, Notes on a schema for stories, in *Representation and Understanding* (Elsevier, New York, 1975), pp. 211–236.
- [11] Roger C. Schank and Robert P. Abelson, *Scripts, Plans, Goals, and Understanding: An Inquiry Into Human Knowledge Structures* (Psychology Press, New York, 2013).
- [12] Perry W. Thorndyke and Frank R. Yekovich, A critique of schema-based theories of human story memory, *Poetics* **9**, 23 (1980).
- [13] Walter Kintsch, *Comprehension: A Paradigm for Cognition* (Cambridge University Press, Cambridge, England, 1998).
- [14] Michael J. Kahana, Computational models of memory search, *Annu. Rev. Psychol.* **71**, 107 (2020).
- [15] Michelangelo Naim, Mikhail Katkov, Sandro Romani, and Misha Tsodyks, Fundamental law of memory recall, *Phys. Rev. Lett.* **124**, 018101 (2020).
- [16] A. Georgiou, T. Can, M. Katkov, and M. Tsodyks, Large-scale study of human memory for meaningful narratives, *Learn. Mem.* **32**, a054043 (2025).
- [17] Tom Trabasso and Linda L. Sperry, Causal relatedness and importance of story events, *J. Mem. Lang.* **24**, 595 (1985).
- [18] Michael J. Kahana, Associative retrieval processes in free recall, *Mem. Cogn.* **24**, 103 (1996).
- [19] Endel Tulving, Subjective organization in free recall of “unrelated” words, *Psychol. Rev.* **69**, 344 (1962).
- [20] Marc W. Howard and Michael J. Kahana, Contextual variability and serial position effects in free recall, *J. Exp. Psychol.* **25**, 923 (1999).
- [21] William Labov, *The Language of Life and Death: The Transformation of Experience in Oral Narrative* (Cambridge University Press, Cambridge, England, 2013).
- [22] W. Kintsch, Toward a model of text comprehension and production, *Psychol. Rev.* **85**, 363 (1978).
- [23] Jean M. Mandler and Nancy S. Johnson, Remembrance of things parsed: Story structure and recall, *Cogn. Psychol.* **9**, 111 (1977).
- [24] Morton Ann Gernsbacher, *Language Comprehension as Structure Building* (Psychology Press, New York, 2013).
- [25] Roy Bar-Haim, Lilach Eden, Roni Friedman, Yoav Kantor, Dan Lahav, and Noam Slonim, From arguments to key points: Towards automatic argument summarization, in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, edited by D. Jurafsky, J. Chai, N. Schluter, and J. Tetreault (Association for Computational Linguistics, 2020), pp. 4029–4039.
- [26] Arthur C. Graesser, Keith K. Willis, and Rolf A. Zwaan, Discourse comprehension, *Annu. Rev. Psychol.* **48**, 163 (1997).
- [27] T. A. van Dijk and W. Kintsch, *Strategies of Discourse Comprehension*, Monograph Series (Academic Press, New York, 1983).
- [28] Robert F. Lorch and Elizabeth P. Lorch, Topic structure representation and text recall, *J. Educ. Psychol.* **77**, 137 (1985).
- [29] Arie Cattan, Lilach Eden, Yoav Kantor, and Roy Bar-Haim, From key points to key point hierarchy: Structured and expressive opinion summarization, in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Toronto, Canada, edited by Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki (Association for Computational Linguistics, Toronto, 2023), pp. 912–928.
- [30] Gabriel A. Radvansky and Jeffrey M. Zacks, Event boundaries in memory and cognition, *Curr. Opin. Behav. Sci.* **17**, 133 (2017).
- [31] Jeffrey M. Zacks, Todd S. Braver, Margaret A. Sheridan, David I. Donaldson, Abraham Z. Snyder, John M. Ollinger, Randy L. Buckner, and Marcus E. Raichle, Human brain activity time-locked to perceptual event boundaries, *Nat. Neurosci.* **4**, 651 (2001).
- [32] Aya Ben-Yakov and Richard N. Henson, The hippocampal film editor: Sensitivity and specificity to event boundaries in continuous experience, *J. Neurosci.* **38**, 10057 (2018).
- [33] Kristjan Kalm, Matthew H. Davis, and Dennis Norris, Neural mechanisms underlying the grouping effect in short-term memory, *Hum. Brain Mapp.* **33**, 1634 (2012).
- [34] Christopher Baldassano, Janice Chen, Asieh Zadbood, Jonathan W. Pillow, Uri Hasson, and Kenneth A. Norman, Discovering event structure in continuous narrative perception and memory, *Neuron* **95**, 709 (2017).
- [35] Jie Zheng, Andrea G. P. Schjetnan, Mar Yebra, Bernard A. Gomes, Clayton P. Mosher, Suneil K. Kalia, Taufik A. Valiante, Adam N. Mamelak, Gabriel Kreiman, and Ueli Rutishauser, Neurons detect cognitive boundaries to structure episodic memories in humans, *Nat. Neurosci.* **25**, 358 (2022).
- [36] Benjamin N. Colby, A partial grammar of eskimo folktales 1, *Am. Anthropol.* **75**, 645 (1973).
- [37] William Labov and Joshua Waletzky, Narrative analysis: Oral versions of personal experience, in *Essays on the Verbal and Visual Arts: Proceedings of the 1966 Annual Spring Meeting of the American Ethnological Society*, edited by June Helm (1966), pp. 12–44.
- [38] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/g1cz-wk11> for additional information about the analytical derivations and the experimental data analysis, which includes Refs. [39–41].
- [39] Herbert A. Simon, How big is a chunk? By combining data from several experiments, a basic human memory unit can be identified and measured, *Science* **183**, 482 (1974).
- [40] Mikhail Katkov, Michelangelo Naim, Antonios Georgiou, and Misha Tsodyks, Mathematical models of human memory, *J. Math. Phys. (N.Y.)* **63**, 073303 (2022).
- [41] Helmut Hasse, Ein summierungsverfahren für die riemannsche ζ -reihe, *Math. Z.* **32**, 458 (1930).

- [42] Nelson Cowan, The magical number 4 in short-term memory: A reconsideration of mental storage capacity, *Behav. Brain Sci.* **24**, 87 (2001).
- [43] P Flajolet, *Analytic Combinatorics* (Cambridge University Press, Cambridge, England, 2009).
- [44] All the data analyzed here are available online at https://github.com/zhongweishun/RandomTreeModel_PRL, which also includes the code to reproduce all the figures used in this Letter.
- [45] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat *et al.*, GPT-4 technical report, [arXiv:2303.08774](https://arxiv.org/abs/2303.08774).
- [46] B. Johnstone, *Stories, Community, and Place: Narratives from Middle America* (Indiana University Press, Bloomington, 1990).
- [47] <https://www.prolific.com>
- [48] Aixin Liu, Bei Feng, Bing Xue, Bingxuan Wang, Bochao Wu, Chengda Lu, Chenggang Zhao, Chengqi Deng, Chenyu Zhang, Chong Ruan *et al.*, Deepseek-v3 technical report, [arXiv:2412.19437](https://arxiv.org/abs/2412.19437).